

PRACTICE PERSPECTIVES

V.G. Viertel
L.S. Babiarz
M. Carone
J.S. Lewin
D.M. Yousem



Quality Control in Neuroradiology: Impact of Trainees on Discrepancy Rates

BACKGROUND AND PURPOSE: Prior studies have found a 2%–8% clinically significant error rate in radiology practice. We compared discrepancy rates of studies interpreted by subspecialty-trained neuroradiologists working with and without trainees.

MATERIALS AND METHODS: Subspecialty-trained neuroradiologists reviewed 2162 studies during 41 months. Discrepancies between the original and “second opinion” reports were scored: 1, no change; 2, clinically insignificant detection discrepancy; 3, clinically insignificant interpretation discrepancy; 4, clinically significant detection discrepancy; and 5, clinically significant interpretation discrepancy. Faculty alone versus faculty and trainee discrepancy rates were calculated.

RESULTS: In 87.6% (1894/2162), there were no discrepancies with the original report. The neuro-radiology division had a 1.8% (39/2162; 95% CI, 1.3%–2.5%) rate of clinically significant discrepancies. In cases reviewed solely by faculty neuroradiologists (16.2% = 350/2162 of the total), the rate of discrepancy was 1.7% (6/350). With fellows (1232/2162, 57.0% of total) and residents (580/2162, 26.8% of total), the rates of discrepancy were 1.6% (20/1232) and 2.2% (13/580), respectively. The odds of a discrepant result were 26% greater (OR = 1.26; 95% CI, 0.38–4.20) when reading with a resident and 8% less (OR = 0.92; 95% CI, 0.35–2.44) when reading with a fellow than when reading alone.

CONCLUSIONS: There was a 1.8% rate of clinically significant detection or interpretation discrepancy among academic neuroradiologists. The difference in the discrepancy rates between faculty only (1.7%), fellows and faculty (1.6%), and residents and faculty (2.2%) was not statistically significant but showed a trend indicating that reading with a resident increased the odds of a discrepant result.

ABBREVIATIONS: CI = confidence interval; OR = odds ratio; QA = quality assurance; SD = standard deviation

To improve the level of patient care that is delivered, one must investigate variables that pervade the patient care process for their impact on the care received. Level of care has been made quantifiable by QA programs, including peer review. Previous publications have found error rates for diagnostic imaging to be between 2.0% and 7.7%.^{1–9} More than 60 years ago, L.H. Garland found that in an environment in which disease prevalence reached 100%, the radiologic error rate was approximately 30%. He then hypothesized that the error rate in a typical radiology practice, one in which 9 of 10 films had negative findings, would be approximately 5%.⁵ His hypothesis was proved many times over, for example when Filippi et al³ examined discrepancy rates between radiology residents and attending neuroradiologists in interpreting emergent neuroradiology MR imaging studies and found a major discrepancy rate of 4.2%. More recently, Zan et al⁹ sampled 4534 neuroradiology cases with an outside report for comparison and found that 347 (7.7%) had clinically significant discrepancies between the outside study and the interpretation of subspecialty-trained neuroradiologists. Using a similar grading system, Babiarz and Yousem¹ performed a study in which 1000 studies were internally reviewed, and they

found a significant discrepancy rate of 2.0% among subspecialty-trained neuroradiologists at a major university hospital.

The purpose of this study was to investigate 1 of the variables affecting the quality of patient care provided by a neuro-radiology service: trainee participation in the initial study review process. The authors hypothesized that having 2 readers review a case would produce more accurate results, even if the second reviewer was a trainee. The authors predicted that when a sample of neuroradiology studies read by faculty alone, faculty with fellows, and faculty with residents was reviewed for accuracy by a second neuroradiology faculty member, the cases initially interpreted solely by faculty members would have the highest discrepancy rates. This article, therefore, compares the discrepancy rates of studies interpreted by subspecialty-trained neuroradiologists working with and without trainees.

Materials and Methods

Data Collection

In accordance with the Health Insurance Portability and Accountability Act, our institutional review board reviewed and approved the protocol for this retrospective study and waived the requirement for informed consent.

The study was conducted as described by Babiarz and Yousem.¹ During 41 months (January 1, 2009, to July 7, 2011), as part of a QA initiative in our department, staff neuroradiologists reviewed previously read neuroradiology studies. For the first 2 current studies of the day that had prior examinations, each neuroradiologist was instructed to review and grade the most recent comparison study and its

Received August 19, 2011; accepted after revision September 20.

From the Russell H. Morgan Department of Radiology and Radiological Sciences (V.G.V., L.S.B., J.S.L., D.M.Y.), The Johns Hopkins Medical Institutions, Baltimore, Maryland; and Division of Biostatistics (M.C.), School of Public Health, University of California, Berkeley, Berkeley, California.

Please address correspondence to D.M. Yousem, MD, The Johns Hopkins Medical Institutions, 600 N Wolfe St, Phipps B-100F, Baltimore, MD 21287; e-mail: dyousem1@jhmi.edu
<http://dx.doi.org/10.3174/ajnr.A2933>

Table 1: Five-point scale used in scoring discrepancies between the original report and the second opinion reading of a study

Point Scale	Verbal Descriptor
1	No change
2	Detection discrepancy, not clinically significant
3	Interpretation discrepancy, not clinically significant
4	Detection discrepancy, clinically significant
5	Interpretation discrepancy, clinically significant

report. The second reviewer graded the original reports on the basis of their own opinion of the findings and interpretation for discrepancies according to a previously validated 5-point rating scale.^{1,9} The 5-point scale allowed the following categories 1) no change in the reading, 2) finding of a clinically insignificant detection discrepancy (eg, missed case of mild chronic sinusitis), 3) finding of a clinically insignificant interpretation discrepancy (eg, interpretation of an oligodendroglioma as an astrocytoma), 4) finding of a clinically significant detection discrepancy (eg, a missed tumor), and 5) finding of a clinically significant interpretation discrepancy (eg, interpreting a tumor as a stroke) (Table 1). In addition to the discrepancy score, staff also noted the type of imaging study of each examined case. In the event of a discrepancy (grades 2–5), staff recorded the source and nature of the discrepancy.

From the pool of the reviewed cases, 200 cases were randomly selected, and their reports were analyzed for predicting practice disease prevalence.

The data from each QA case sheet were input into a spreadsheet, and we recorded the following: patient medical record number, date of the original study, date of the second review, type of examination, technique and original contributor, reviewer, score, and comments. The electronic patient record was queried to determine whether the original faculty contributor was the sole contributor on the original report or if a trainee's name also appeared as a contributor. Diagnostic radiology resident and neuroradiology fellow trainees were separately designated.

Nineteen subspecialty-certified or subspecialty-certification-eligible neuroradiologists assigned to read current cases reviewed cases and reports originally read by 23 subspecialty-certified or subspecialty-certification-eligible neuroradiologists. Four neuroradiologists were no longer within the division and represent former faculty who had read cases. A neuroradiologist is subspecialty-certified or subspecialty-certification-eligible following completion of a 1-year Accreditation Council for Graduate Medical Education–accredited diagnostic neuroradiology fellowship. Neuroradiologists must practice neuroradiology for 1 year after their fellowship year and then pass a 4-hour cognitive test of neuroradiology knowledge, proctored by members of the American Board of Radiology, to become subspecialty-certified.

Each neuroradiologist was responsible for covering the clinical service an average of 3–4 days a week. On some services, there may not have been 2 cases to review that had comparison studies because of low volumes (eg, myelography service, teleradiology service, and so forth). The 19 neuroradiologist reviewers were not allowed to review their own previous radiology reports and were instructed to skip such cases. The readers ranged in experience from 1 to 29 years' post-neuroradiology fellowship training.

For validating the scoring system, all studies with clinically significant discrepancies (scores of 4 and 5) were assessed by a third independent expert reviewer with 23 years of experience in neuroradiol-

Table 2: Final distribution of discrepancy scores

Score	No. of Studies	% Total
1	1894	87.6
2	168	7.8
3	61	2.8
4	32	1.5
5	7	0.3
1, 2, and 3	2123	98.2
4 and 5	39	1.8
Total	2162	100.0

ogy (D.M.Y.) who had previously published data using the 5-point scoring system.^{1,8} Because D.M.Y. was 1 of the 23 neuroradiologists whose reports were reviewed in this study, no adjudicating was done on the single discrepant case that he had originally read.

Statistical Analysis

Total counts of all scores and their relative percentages were tabulated for each staff member as well as for the whole neuroradiology division. The discrepancy rate was defined as the proportion of studies resulting in clinically significant findings. For the purpose of calculating the overall and individual discrepancy rates, scores were dichotomized as being either above 3 (clinically significant findings) or not greater than 3 (clinically insignificant findings). Basic descriptive statistics, including the range, mean, and SD, were computed for observed counts. Discrepancy rates were calculated for each imaging technique (CT versus MR imaging) and body region imaged (brain/head and neck versus spine versus bony structures). Discrepancy rates were also calculated by type of contributors, including faculty only, faculty and neuroradiology fellow, and faculty and resident. ORs for discrepant results comparing studies with faculty interpretation alone with studies including residents and fellows were computed. CIs for discrepancy rates and ORs were constructed by inverting asymptotically normal test statistics obtained from appropriate logistic regression models. These models were fitted by using generalized estimating equations to account for the potential correlation between studies conducted by the same contributor. All tests of hypotheses were 2-sided and conducted at significance level of .05.

Results

The 2162 neuroradiology studies originally read by 23 neuroradiologists (present and past) were reread by the 19 neuroradiologists on staff at the time of their review. On average, 94.0 studies were reviewed for each of the 23 original readers (range, 1–258; SD, 83.8).

The final scores of the cases following adjudication yielded 2123 scores of 1, 2, or 3 (the same or with insignificant discrepancies) and 39 scores of 4 or 5 (significant discrepancies) (Table 2). In 87.6% of cases (1894 of 2162; 95% CI, 85.7%–89.3%), there were no changes recommended to the report (score = 1).

Of the cases read by faculty only, which made up 16.2% of the total cases (350 of 2162), 6 had a score of 4 or 5 (1.7%). Of the cases read by faculty with fellows, which made up 57.0% of the total cases (1232 out of 2162), 20 had a score of 4 or 5 (1.6%). There was a total of 54 fellows who contributed to reports. The average fellow read 22.5 cases (range, 1–85; SD, 24.7). The personal discrepancy rates of fellows ranged from 0.0% to 10.0%. As for cases read by faculty and residents, which made up the remaining 26.8% (580 of 2162), 13 had a

Table 3: Discrepant cases by disease type (scores of 4 or 5)		
Discrepancy Category	Discrepancy Score (4 or 5)	Imaging Study Type
Artifacts	4	CT brain
Artifacts	5	CT brain
Congenital	5	CT brain
Congenital	5	CT brain
Degenerative	4	MRI spine
Degenerative	4	CT spine
Infection	4	CT maxillofacial
Neoplasm	4	CT brain
Neoplasm	4	CT brain
Neoplasm	4	MRI brain
Neoplasm	4	MRI brain
Neoplasm	4	MRI neck
Neoplasm	4	MR spine
Neoplasm	4	CT neck
Neoplasm	4	CT neck
Neoplasm	4	MRI brain
Neoplasm	4	MRI brain
Neoplasm	4	CT head
Neoplasm	4	CT head
Neoplasm	4	CT neck
Neoplasm	4	MR brain
Neoplasm	5	CT brain
Neoplasm	5	MRI brain
Neoplasm	5	MRI brain
Neoplasm	5	MRI neck
Trauma	4	MRI brain
Trauma	4	CT head
Vascular	4	CT angiography
Vascular	4	MRI brain
Vascular	4	MRI brain
Vascular	4	MRI brain
Vascular	4	MRI brain
Vascular	4	CT brain
Vascular	4	CT brain
Vascular	4	MRI brain
Vascular	4	MRI venography
Vascular	4	CT head
Vascular	4	MRI brain
Vascular	5	MRI brain

score of 4 or 5, yielding a discrepancy rate of 2.2% (13 of 580), the highest recorded.

There was a total of 79 residents who contributed to reports. The average resident read 7.5 cases (range, 1–35; SD, 7.5). The personal discrepancy rates of residents ranged from 0.0% to 100.0%. The differences between the discrepancy rates for faculty alone, with fellows, and with residents were not statistically significant. For a faculty member, the odds of a discrepant result were 26% greater (OR = 1.26; 95% CI, 0.38–4.20) when reading with a resident and 8% less (OR = 0.92; 95% CI, 0.35–2.44) when reading with a fellow than when reading alone.

Of the 39 cases with scores of 4 or 5, nineteen were CTs (13 of the brain, 3 of the neck, 1 of the maxillofacial bones, 1 of the spine, 1 CT angiogram) and 20 were MR images (15 of the brain, 2 of the neck, 2 of the spine, and 1 MR venogram). Of the 39 “misses,” 12 cases were classified as vascular; 18, as neoplasms; 2, as congenital; 2, as artifacts; 2, as degenerative; 2, as trauma; and 1, as infection (Table 3).

Of the 200 randomly selected reviewed cases, 20 (10%) were studies with normal findings (no abnormalities) and 180

Table 4: Discrepancy by type of examination				
Modality/Type of Exam	Total	% of Total	No. of Discrepant Cases	Discrepancy Rate of the Specific Types of Exams
CT				
Brain/head & neck	675	31.2	17	2.5
Spine	43	2.0	1	2.3
Bony face & neck structures	158	7.3	1	.6
Total	879	40.5	19	2.2
MR imaging				
Brain/head & neck	1155	53.4	18	1.6
Spine	119	5.5	2	1.7
Face, pituitary, skull base	6	0.3	0	0.0
fMRI	1	.05	0	0.0
Total	1281	59.3	20	1.6
Other				
Angiogram	1	.05	0	0.0
Radiograph	1	.05	0	0.0
Myelogram	3	.1	0	0.0
Overall	2162	100	39	1.8

(90%) had positive radiographic findings. There were 74 (37%) neoplastic, 33 (16.5%) vascular (stroke, aneurysms), 17 (8.5%) iatrogenic (tumor resection, ventricular shunt placement), 19 (9.5%) infectious/inflammatory (encephalitis, abscess, demyelinating disease), 15 (7.5%) traumatic, 11 (5.5%) congenital, 5 (2.5%) degenerative (spine), and 6 (3%) miscellaneous disease cases. Applying these ratios to the entire 2162 studies yielded clinically significant detection or interpretation discrepancy rates of 2 of 119 (1.7%) congenital abnormalities, 1 of 205 (0.5%) infectious abnormalities, 2 of 54 (3.7%) degenerative abnormalities, 2 of 162 (1.2%) traumatic abnormalities, 12 of 356 (3.4%) vascular abnormalities, and 18 of 800 (2.3%) neoplastic abnormalities. There was no statistically significant difference in clinically significant discrepancy rates among these categories.

For all CT studies, the rate of clinically significant detection or interpretation discrepancies (scores of 4 or 5) was 2.2% (19 of 876; 95% CI, 1.3%–3.6%) and for all MR imaging studies, it was 1.6% (20 of 1281; 95% CI, 1.0%–2.4%). There were no statistically significant differences between CT and MR imaging discrepancy rates ($P = .33$). For all images of the brain/head and neck, the discrepancy rate was 1.9% (35 in 1830; 95% CI, 1.4%–2.7%); for all images of the spine, it was 1.9% (3 in 162; 95% CI, 0.6%–5.6%); and for the images of the facial bones (including maxillofacial bones, orbits, sinuses, and temporal bone), the rate was 0.6% (1 in 160; 95% CI, 0.1%–4.5%; Table 4).

Overall, the neuroradiology service had a 1.8% (39 in 2162; 95% CI, 1.3%–2.5%) rate of clinically significant detection or interpretation discrepancies (scores of 4 or 5) when images were reread (Table 2). Although 23 neuroradiologists had dictated studies that were reviewed, only 16 of the 23 had ≥ 10 studies as part of their personal evaluation. For these 16 neuroradiologists who dictated the ≥ 10 original reports, the discrepancy rate ranged between 0.4% and 5.7% (mean = 1.8%; SD = 1.7%). Table 5 provides individual discrepancy rates for all 23 neuroradiologists.

Table 5: Individual discrepancy rate

Reviewer	Total Studies Read	Scores of 4 and 5	Discrepancy Rate
A	1	0	0.0%
B	1	0	0.0%
C	2	0	0.0%
D	3	0	0.0%
E	4	0	0.0%
F	5	0	0.0%
G	8	0	0.0%
H	258	1	0.4%
I	156	1	0.6%
J	175	1	0.6%
K	175	2	1.1%
L	183	2	1.1%
M	238	3	1.3%
N	196	4	2.0%
O	140	3	2.1%
P	40	1	2.5%
Q	141	4	2.8%
R	98	3	3.1%
S	31	1	3.2%
T	63	2	3.2%
U	81	3	3.7%
V	116	5	4.6%
W	53	3	5.7%
Overall	2162	39	1.8%

Discussion

Research on the effect of factors impacting the image interpretation process is essential for the creation of an environment that is conducive to the highest possible level of patient care. In the past, many variables have been analyzed for their effect on the quality of service, such as distractions, availability of clinical history, availability of previous reports, duration of search time, reading room environment, and trainee participation.^{10–13}

Although our study focuses primarily on the effect of trainees on discrepancy rates, previous studies have analyzed similar variables, particularly how accurate residents are in addition to how their residency year affects their accuracy.^{7,14} For example, Seltzer et al¹⁵ found that the need for correction of reports made by residents decreased significantly between the first and second/third years of residency. When Filippi et al³ looked specifically at emergency neuroradiology MR imaging studies and year of residency, the percentages ranged from 10.9% with first-year residents to 4.7% with second-year residents. Cooper et al¹⁴ had similar results when they found that fourth-year residents had a statistically significantly lower rate of error, compared with doctors earlier in their residency. Although our study does not compare year of residency, it does compare a similar variable: level of expertise (ie, between residents, fellows, and faculty). It is the goal of residency and fellowship programs to train individuals with the hope of producing better health care professionals; thus, it is important to periodically ensure that such training is effective and improves health care as a whole.

The participation of trainees in the image-interpretation process, most commonly through review of studies before faculty interpretation, has many benefits. It is a critical component in the development of trainee proficiency and the achievement of eventual autonomy. In addition, previous

studies have shown that multiple readers may be beneficial for interpretation accuracy and, therefore, patient care. Sickles et al¹⁶ conducted a study analyzing mammography reports and found a significantly higher performance by radiologists using a multiple rather than a single reading system of interpretation. Some studies in mammography have followed the effect of a quadruple reading system, while others have varied as to whether the second reader was blinded to the first reader's review.^{17,18}

In this study, we found a 1.8% (39/2162; 95% CI, 1.2%–2.4%) rate of clinically significant detection or interpretation discrepancy between the original report and the second opinion review of 2162 neuroradiology examinations. The odds of a discrepant result when a faculty member was reading with a resident were 26% (OR = 1.26; 95% CI, 0.38–4.20) greater than when he or she was working alone. This rate was 34% greater than when the faculty member was reading with a fellow. The odds of a discrepant result when a faculty member was reading with a fellow were 8.0% (OR = 0.92; 95% CI, 0.35–2.44) less than those when the faculty member was reading alone. These differences do not achieve statistical significance, though they do reflect a range of discrepancy rates that call to question the impact of trainee participation in the reading process. Our rates of discrepancy of 1.8% are less than those studies cited above that refer to residents' readings that are checked by faculty.^{3,14,15} The homogeneity of readings and lower rates are likely due to the advanced state of training of all of the neuroradiology faculty and the similarities in case mix represented.

The focus of our study was on the impact of resident and fellow participation in the reading of neuroradiology scans and the subsequent discrepancy rates. The implications of our study are that a trend can be seen in a higher rate of discrepancies when residents are involved. We cannot determine the accuracy rates of the residents in this study, which would be useful in making suppositions from the data. However, 1 explanation may be that the faculty members had unwarranted confidence in the residents' interpretation and did not apply the same level of scrutiny to the findings as they would reading the case themselves. Was there a sense of complacency in approving residents' reports? If so, that confidence may not have been as well-placed as when one reviews a board-certified neuroradiology fellow. In our department, residents and fellows dictate cases on their own, and their completed reports are then revised and/or approved by the faculty. Also, most studies read at our institution (1812 of 2162, or 83.8% of cases included in this study) are read by faculty and residents/fellows.

There are a number of study limitations. First, for a study that spanned >41 months with the intention of 2 studies reviewed per day per neuroradiologist on clinical service, the total number of studies was much lower than expected. This resulted from a lower compliance rate than desired. Next, the scoring system, which has been validated by prior studies,^{1,9} was somewhat arbitrary, and its use may have led to misclassification or a failure to capture some of the differences between the original reports and the reviews of these reports. In addition, the design is susceptible to confirmation bias. Because reviewers knew how the disease progressed in the follow-up imaging, they may have had an easier time detecting

the more subtle cues that presented themselves in the original images.

Another bias that limited this study was disease-selection bias. Because we only included cases for which follow-up studies were being done, we likely altered the distribution of diseases seen and types of examinations reviewed. There is a wide range in the number of cases reviewed per faculty member largely based on the percentage of time on the clinical schedule. This means that some “funded” faculty members are under-represented in the sample. What we are reporting in this study is discrepancy rates, and they are not the equivalent of accuracy rates. Unfortunately, because of inadequate/incomplete follow-up for determination of definitive diagnosis in most cases, accuracy rates cannot be computed. Finally, another important limitation is the potential bias from having 1 expert reviewer in charge of final adjudication of all (except for his own) clinically significant discrepancies.

There are a number of ways in which the QA program can be improved to minimize these limitations. For example, to improve care for included patients, cases could be reviewed by a second neuroradiologist immediately following their reading, before releasing the report to the referring physician. In addition, the design of future studies with an assessment of accuracy instead of just discrepancy would also be beneficial.

Conclusions

At a university hospital, we found a 1.7% rate of clinically significant detection or interpretation discrepancy between the original dictations and the second-opinion reviews of neuroradiology cases among faculty neuroradiologists working alone, a 1.6% rate of discrepancy among faculty neuroradiologists working with fellows, and a 2.2% rate of discrepancy among faculty neuroradiologists working with residents. The overall rate of discrepancy was 1.8%. The results suggest that greater scrutiny may be warranted in reviewing residents’ interpretations of neuroradiology studies. Our results may have limited generalizability because we only reviewed cases for which follow-up studies were being done and thus potentially introduced disease-selection bias. Also, our study design may have introduced confirmation bias by allowing the reviewers to see how the cases evolved with time. Our study is an example of 1 step in the practice quality-improvement process that we hope will serve as a baseline and outcome metric as we intervene to reduce variability in interpretation quality. Accuracy rates must also be addressed for the best patient care.

Acknowledgments

We thank Tom Deluca and Gloria Vila for IT Radiology Information System search assistance.

Disclosures: Jonathan S. Lewin—*UNRELATED: Board Membership: H2MI LLC,* Comments: I serve on the Board of Managers for H2MI, a company with affiliation to Johns Hopkins Medicine and Harris Corporation (this company has no relevance to this article); Grants/Grants Pending: Research Agreements,* Patents (planned, pending, or issued): various, Comments: I am coinventor on 19 issued and 3 pending US patents and a number of international patents, unrelated to the this article; Stock/Stock Options: Interventional Imaging, Inc., Comments: I am a cofounder and shareholder of Interventional Imaging, Inc, which does not have relevance to this article; Other: BAM Laboratories, Comments: I hold stock options in BAM laboratories, a technology firm unrelated to the current article.*Money paid to the institution.*

References

1. Babiarz LS, Yousem DM. **Quality control in neuroradiology: discrepancies in image interpretation among academic neuroradiologists.** *AJNR Am J Neuroradiol* 2012;33:37–42
2. Borgstede JP, Lewis RS, Bhargavan M, et al. **RADPEER quality assurance program: a multifacility study of interpretive disagreement rates.** *J Am Coll Radiol* 2004;1:59–65
3. Filippi CG, Schneider B, Burbank HN, et al. **Discrepancy rates of radiology resident interpretations of on-call neuroradiology MR imaging studies.** *Radiology* 2008;249:972–79
4. Garland LH. **On the scientific evaluation of diagnostic procedures.** *Radiology* 1949;52:309–28
5. Garland LH. **Studies on the accuracy of diagnostic procedures.** *Am J Roentgenol Radium Ther Nucl Med* 1959;82:25–38
6. Robinson PJ, Wilson D, Coral A, et al. **Variation between experienced observers in the interpretation of accident and emergency radiographs.** *Br J Radiol* 1999;72:323–30
7. Siegle RL, Baram EM, Reuter SR, et al. **Rates of disagreement in imaging interpretation in a group of community hospitals.** *Acad Radiol* 1998;5:148–54
8. Soffa DJ, Lewis RS, Sunshine JH, et al. **Disagreement in interpretation: a method for the development of benchmarks for quality assurance in imaging.** *J Am Coll Radiol* 2004;1:212–17
9. Zan E, Yousem DM, Carone M, et al. **Second-opinion consultations in neuro-radiology.** *Radiology* 2010;255:135–41
10. Berlin L. **Comparing new radiographs with those obtained previously.** *AJR Am J Roentgenol* 1999;172:3–6
11. Berlin L. **Accuracy of diagnostic procedures: has it improved over the past five decades?** *AJR Am J Roentgenol* 2007;188:1173–78
12. Loy CT, Irwig L. **Accuracy of diagnostic tests read with and without clinical information: a systematic review.** *JAMA* 2004;292:1602–09
13. Oestmann JW, Greene R, Kushner DC, et al. **Lung lesions: correlation between viewing time and detection.** *Radiology* 1988;166:451–53
14. Cooper VF, Goodhart LA, Nemcek AA Jr, et al. **Radiology resident interpretations of on-call imaging studies: the incidence of major discrepancies.** *Acad Radiol* 2008;15:1198–204
15. Seltzer SE, Hessel SJ, Herman PG, et al. **Resident film interpretation and staff review.** *AJR Am J Roentgenol* 1981;137:129–33
16. Sickles EA, Wolverton DE, Dee KE. **Performance parameters for screening and diagnostic mammography: specialist and general radiologists.** *Radiology* 2002;224:861–69
17. Duijm LE, Groenewoud JH, Fracheboud J, et al. **Additional double reading of screening mammograms by radiologic technologists: impact on screening performance parameters.** *J Natl Cancer Inst* 2007;99:1162–70
18. Mahgerefteh S, Kruskal JB, Yam CS, et al. **Peer review in diagnostic radiology: current state and a vision for the future.** *Radiographics* 2009;29:1221–31